Oxford Cambridge and RSA

GCE

Further Mathematics B (MEI)

Y422/01: Statistics major

Advanced GCE

Mark Scheme for Autumn 2021

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of candidates of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, Cambridge Nationals, Cambridge Technicals, Functional Skills, Key Skills, Entry Level qualifications, NVQs and vocational qualifications in areas such as IT, business, languages, teaching/training, administration and secretarial skills.

It is also responsible for developing new specifications to meet national requirements and the needs of students and teachers. OCR is a not-for-profit organisation; any surplus made is invested back into the establishment to help towards the development of qualifications and support, which keep pace with the changing needs of today's society.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by examiners. It does not indicate the details of the discussions which took place at an examiners' meeting before marking commenced.

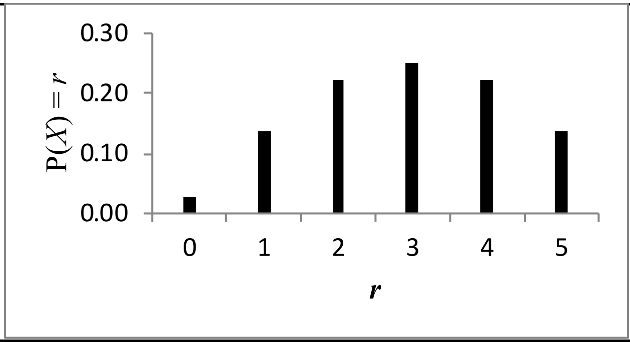
All examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the report on the examination.

Annotations and abbreviations

Annotation in scoris	Meaning
✓ and ✗	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0, 1
E	Explanation mark 1
SC	Special case
^	Omission sign
MR	Misread
BP	Blank page
Highlighting	
Other abbreviations in mark scheme	**Meaning**
E1	Mark for explaining a result or establishing a given result
dep*	Mark dependent on a previous mark, indicated by *. The * may be omitted if only previous M mark.
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working
AG	Answer given
awrt	Anything which rounds to
BC	By Calculator
DR	This indicates that the instruction **In this question you must show detailed reasoning** appears in the question.

Question			Answer	Marks	AOs	Guidance	
1	(a)		34.711 ± 1.96 $\times \frac{1.53}{\sqrt{50}}$ $= 34.711 \pm 0.424$ or $(34.287, 35.135)$	**B1** **M1** **M1** **A1** **[4]**	**1.1** **3.3** **1.1** **3.4**	Allow 34.29 to 35.13 or 35.14	
1	(b)		50 is a sufficiently large sample to apply the CLT which states that for large samples the distribution of the sample mean is approximately Normal	**B1*** ***B1** **[2]**	**2.2b** **2.4**	For mention of central limit theorem For full statement (including CLT)	No credit if CLT not mentioned

Question			Answer	Marks	AOs	Guidance	
2	(a)		$P(X=0)=\frac{6}{6}\times\frac{1}{6}\times\frac{1}{6}$ $=\frac{1}{36}$	**M1** **A1** **[2]**	**3.1a** **1.1**	 **AG**	Allow M1 for $\frac{1}{6}\times\frac{1}{6}=\frac{1}{36}$
2	(b)			**B1** **B1** **[2]**	**1.1** **1.1**	For heights For axes and labels	Roughly correct but must have linear scale Do not allow just P on vertical axis
2	(c)		The distribution has (slight) negative skew	**B1** **[1]**	**1.1**	Allow 'roughly symmetrical' or 'unimodal'	Not 'Normal distribution'
2	(d)		**DR** $E(X)=0\times\frac{1}{36}+1\times\frac{5}{36}+2\times\frac{2}{9}+3\times\frac{1}{4}+4\times\frac{2}{9}+5\times\frac{5}{36}$ $=\frac{105}{36}=\frac{35}{12}=2.9166\ldots$ $E(X^2)=0^2\times\frac{1}{36}+1^2\times\frac{5}{36}+2^2\times\frac{2}{9}+3^2\times\frac{1}{4}+4^2\times\frac{2}{9}+5^2\times\frac{5}{36}$ $=\frac{371}{36}=10.3055\ldots$ $\text{Var}(X)=10.3055\ldots-(2.9166\ldots)^2$ $=\frac{259}{144}=1.80\quad(1.7986\ldots)$	**M1** **A1** **M1** **M1** **A1** **[5]**	**1.1a** **1.1** **1.1** **1.2** **1.1**	 Allow fraction or decimal form	
2	(e)		$\text{Variance}=30^2\times1.7986\ldots=1619\ (\text{pence}^2)$	**B1** **[1]**	**1.1**		
2	(f)		Average amount received = 30 × 2.916… = 87.5 $k-87.5=12.5\Rightarrow k=100$	**B1** **B1** **[2]**	**3.1a** **1.1**		

Question			Answer	Marks	AOs	Guidance	
3	(a)		Using B(50, 0.04) $P(X = 2) = 0.276$	M1 A1 [2]	3.3 1.1	 BC	
3	(b)		$0.96^9 \times 0.04 = 0.0277$	B1 [1]	1.1		Allow 0.028
3	(c)		$0.96^{20} = 0.442$	B1 [1]	1.1		
3	(d)		Expected value for one misunderstood $= \frac{1}{0.04} = 25$ Because geometric For 3 misunderstood expected number = 25 + 25 + 25 = 75	B1 E1 E1 [3]	2.1 2.4 1.1		Must quote probabilities to get full marks
3	(e)		Require P(2 misunderstood in first 59) × 0.04 so using B(59, 0.04) gives $P(X = 2) = 0.267$ 0.267 × 0.04 = 0.0107	B1 M1 A1 [3]	3.1a 2.2a 1.1	For identifying required probability Use of correct binomial BC	

4	(a)		Nuclei decay randomly and decays are independent with constant probability $\frac{1}{200\,000}$ The number of decays out of 1 000 000 is being counted, so a binomial distribution is appropriate Because $n = 1\,000\,000$ is large and $p = \frac{1}{200\,000}$ is small a Poisson distribution is also appropriate	E1 E1 E1 [3]	2.4 2.4 2.4	For partial explanation of binomial For full explanation For explanation of Poisson	
4	(b)		Po(5) $P(X = 6) = 0.146$ $P(X > 6) = 1 - 0.762 = 0.238$	M1 A1 A1 [3]	3.3 1.1 1.1	 BC BC	
4	(c)		Mean = 10 × 5 = 50 P(at least 60 decays) $= 1 - 0.9077 = 0.0923$	B1 B1 [2]	3.3 1.1	 BC	Allow 0.092

Question			Answer	Marks	AOs	Guidance	
5	(a)		Two A and one B ~ $N(2 \times 3.9 + 7.8,\ 2 \times 0.32^2 + 0.41^2)$	**B1**	**3.3**	For N and mean	Allow if N stated anywhere in answer SOI
			$N(15.6, 0.3729)$	**M1**	**1.1**	For variance	
			$P(\geq 16) = 0.256$ (0.25622…)	**A1**	**3.4**	**BC**	
				[3]			
5	(b)		Four B – one C ~ $N(4\times7.8 - 30.2,\ 4\times0.41^2 + 0.64^2)$	**B1**	**3.3**	For N and mean	Allow -1 for mean
			$N(1, 1.082)$	**M1**	**1.1**	For variance	Allow if N stated anywhere in answer SOI
			P(within 1 unit) = 0.473 (0.47274…)	**A1**	**3.4**	**BC**	
				[3]			
5	(c)		**DR**				
			$H_0: \mu = 30.2$ $H_1: \mu \neq 30.2$	**B1**	**3.3**	Hypotheses in words only must include "population"	
			where μ is the population mean capacitance	**B1**	**1.2**	For definition in context	
			Sample mean = 29.96	**B1**	**1.1**		
			Est. population variance $= \frac{1}{9}\left(8981.0 - \frac{299.6^2}{10}\right)$	**M1**	**1.1**		
			$= 0.5538$	**A1**	**1.1**		Or sd = 0.7442
			Test statistic $= \frac{29.96 - 30.2}{\sqrt{\frac{0.5538}{10}}}$	**M1**	**3.3**	FT their mean and/or sd	
			$= -1.020$	**A1**	**1.1**	BC	
			Refer to t_9	**M1**	**3.4**	No FT if not t_9	
			Critical value (2-tailed) at 5% level is 2.262	**A1**	**1.1**		Or $P(t < -1.020) = 0.1672$
			$-1.020 > -2.262$ so not significant (do not reject H_0)	**M1**	**2.2b**	Or $1.020 < 2.262$	Or $0.1672 > 0.025$
			Insufficient evidence to suggest that the capacitance of the batch is different from 30.2	**E1**	**3.5a**		Answer must be in context
				[11]			

Question			Answer	Marks	AOs	Guidance	
6	(a)		Mean = 1.725	B1	1.1		Or $\frac{345}{200}$
			Variance = 1.768	B1	1.1	Condone 1.759 (using divisor n)	
			The variance is reasonably close to the mean so this does support the suitability of a Poisson model	E1	2.2b		Dep on mean and variance correct
				[3]			
6	(b)		Cell C3 = 0.3106	B1	3.4		Do not allow 0.311
			Cell D3 = 62.1124	B1FT	2.2a	200 × their C3 (62.12 if use 0.3106)	Allow 62.2 from 0.311
			Cell E3 = $\dfrac{(65-62.1224)^2}{62.1224}$	M1FT	1.1a		Must show working to get M1
			= 0.1342	A1	1.1		Allow 0.126 from 62.2
				[4]			
6	(c)		Because otherwise some expected frequencies would be less than 5 so too small for the test to be valid	E1	3.5b	For 'less than 5 so invalid'	
				[1]			
6	(d)		H_0: Poisson model is a good fit H_1: Poisson model is not a good fit	B1	2.5		
			$X^2 = 2.43$	B1FT	1.1	FT Their value of E3	
			Refer to χ_5^2	B1	3.4	For degrees of freedom = 5 soi	
			Critical value at 5% level = 11.07	B1	1.1		
			2.43 < 11.07 so result is not significant	M1	1.1	For comparison with critical value	Allow M1 (not A1) for comparison with any chi squared critical value eg 1.145 or 5.991
			There is insufficient evidence to suggest that the Po(1.7) model is not a good fit.	A1	2.2b	Conclusion in context	
				[6]			

Question			Answer	Marks	AOs	Guidance	
7	(a)		The pairing will eliminate any differences in grip strengths between different people and so will only compare the grip strengths of the dominant and non-dominant hands	**E1** **E1** **[2]**	**2.2b** **2.2b**	Give 1 mark for any valid comment For 2 marks must include pairing	
7	(b)		The parent population of differences must be Normally distributed	**E1** **E1** **[2]**	**1.1** **1.2**	For Normally distributed For full answer including 'differences'	
7	(c)		It does because the confidence interval contains 2	**E1** **[1]**	**3.5a**		
7	(d)	(i)	Sample mean difference = 2.39 $0.45 = 1.96 \times \frac{\text{SD}}{\sqrt{100}}$ Sample SD = 2.30 (2.2959…)	**B1** **M1** **A1** **[3]**	**1.1** **3.1b** **1.1**		
7	(d)	(ii)	The sample must be random since only a random sample enables proper inference about the population to be undertaken	**B1** **B1** **[2]**	**3.2b** **2.4**	Do not allow eg a random sample is less likely to be biased	

Question			Answer	Marks	AOs	Guidance	
8	(a)	(i)	Predicted = 50.5	B1 [1]	1.1		Do not allow answer to more than 2dp
8	(a)	(ii)	Although this point lies within the data (interpolation), the points do not lie too close to the line and the value of r^2 is not too close to 1 so the estimate is only moderately reliable	B1 B1 [2]	2.2a 3.5b	Mention of 1 of the three points Mention of at least 2 points with correct conclusion	
8	(a)	(iii)	Coordinates (47.3, 48.7)	B1 [1]	1.1		
8	(a)	(iv)	This is the point with coordinates which are the means of the x- and y-values respectively	B1 [1]	1.1	Allow 'This is the centroid'	
8	(b)	(i)	The scatter diagram is very roughly elliptical and so the distribution may be bivariate Normal	E1 E1 [2]	3.5a 2.4		
8	(b)	(ii)	$S_{vt} = 3886.53 - \frac{1}{20} \times 80.37 \times 970.86 \quad (= -14.87\ldots)$ $S_{tt} = 324.71 - \frac{1}{20} \times 80.37^2 \quad (= 1.743\ldots)$ $S_{vv} = 47829.24 - \frac{1}{20} \times 970.86^2 \quad (= 700.78\ldots)$ $r = \frac{S_{tv}}{\sqrt{S_{tt}S_{vv}}} = \frac{-14.87}{\sqrt{1.743 \times 700.78}}$ $= -0.4255$	M1 M1 M1 A1 [4]	1.1a 1.1 3.3 1.1	Numerical evaluations are not required at this stage For either S_{tt} or S_{vv} For general form including sq. root BC	
8	(b)	(iii)	H_0: $\rho = 0$, H_1: $\rho < 0$ where ρ is the population pmcc between t and v For $n = 20$, the 5% critical value is 0.3783 Since $\lvert -0.4255 \rvert > 0.3783$ the result is significant, so there is sufficient evidence to reject H_0 There is sufficient evidence at the 5% level to suggest that there is negative correlation between marathon time and VO_{2max}	B1 B1 B1 M1 A1FT [5]	3.3 2.5 3.4 1.1 2.2b	For both hypotheses For defining ρ For correct critical value For comparison and conclusion Allow $-0.4255 < -0.3783$ FT for conclusion in words	Do not allow r in place of ρ Hypotheses in words only get B1 unless population mentioned Answer must be in context

Question			Answer	Marks	AOs	Guidance	
9	(a)		$\mathrm{P}\left(X > \frac{1}{2}n\right) = \dfrac{\frac{1}{2}(n+1)}{2n+1}$	M1 M1	3.1a 1.1	For correct denominator For correct numerator	
			$= \dfrac{n+1}{2(2n+1)}$	A1	1.1		
				[3]			
9	(b)		$(2n + 1)$ values so $\mathrm{Var}(X) = \frac{1}{12}\left[(2n+1)^2 - 1\right]$	M1	3.1a		
			Var of sum of 10 values $= 10 \times \frac{1}{12}\left[(2n+1)^2 - 1\right]$	M1	1.1		Allow M1 for 10× any attempt at variance
			$= \frac{10}{3}n^2 + \frac{10}{3}n$	A1	1.1		
				[3]			

10	(a)		$\mathrm{P}(T \le 56) = \dfrac{104}{500} = 0.208$	B1	1.1		
			$\mathrm{P}(T > 61) = 1 - \dfrac{253}{500} = 0.494$	B1	1.1		
				[2]			
10	(b)		$\mathrm{E}(T) = 25 + 28 + 5 + 3 = 61$	B1	3.1a		
			$\mathrm{Var}(T) = \frac{1}{12} \times 10^2 + \frac{1}{12} \times 6^2 + 4 + 16$	M1	1.1		
			$= \frac{94}{3} \quad (= 31.333)$	A1	1.1		
			$W \sim \mathrm{N}(61,\ 31.333)$ so $\mathrm{P}(W \le 56) = 0.186$	B1	3.3	BC	
			$\mathrm{P}(W > 61) = 0.5$	B1	1.1		
				[5]			
10	(c)		Because the mean is 61 and both the uniform and Normal distributions are symmetrical so you would expect the simulated probability to be very close to 0.5	E1 E1 [2]	2.2b 2.4	For second mark must mention symmetrical	

Question			Answer	Marks	AOs	Guidance	
11	(a)		$F(3)=1 \Rightarrow \int_0^2 ax^2\,dx + \int_2^3 b(3-x)^2\,dx = 1$	M1	3.1a		
			$\Rightarrow \frac{8}{3}a + \frac{1}{3}b = 1$	A1	1.1		
			$E(X)=2 \Rightarrow \int_0^2 ax^3\,dx + \int_2^3 bx(3-x)^2\,dx = 2$	M1	3.1a		
			$\Rightarrow 4a + \frac{3}{4}b = 2$	A1	1.1		
			$a = \frac{1}{8},\ b = 2$	A1	1.1		
				[5]			
11	(b)		$F(2) = \int_0^2 \frac{1}{8}x^2\,dx = \frac{1}{3}$	B1	3.1a		
			$\Rightarrow \int_2^m 2(3-x)^2\,dx = \frac{1}{6}$	M1	2.2a		
			$\Rightarrow -\frac{2}{3}(3-m)^3 + \frac{2}{3} = \frac{1}{6}$				
			$\Rightarrow (3-m)^3 = \frac{3}{4} \Rightarrow m = 2.09 \quad (2.0914\ldots)$	A1	1.1	Or $m = 3 - \sqrt[3]{\frac{3}{4}}$	
				[3]			
11	(c)		Using $N\left(2,\ \frac{0.2}{50}\right)$	M1	3.1a	For use of Normal distribution	
			N(2, 0.004)	M1	1.1a	For correct values	
			Estimate P(Mean < 1.9) = 0.0569	A1	1.1		
				[3]			

OCR (Oxford Cambridge and RSA Examinations)
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA

OCR Customer Contact Centre

Education and Learning
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk